Experiences with the Dawn BG/P Platform for High Fidelity Semiconductor Device Simulations

LLNL Dawn User Forum (DUF)

User Presentation

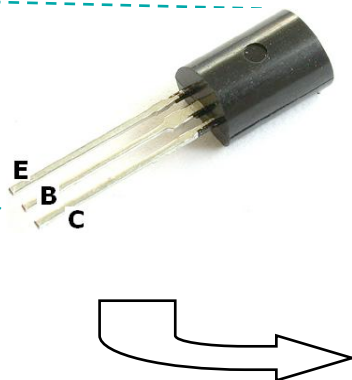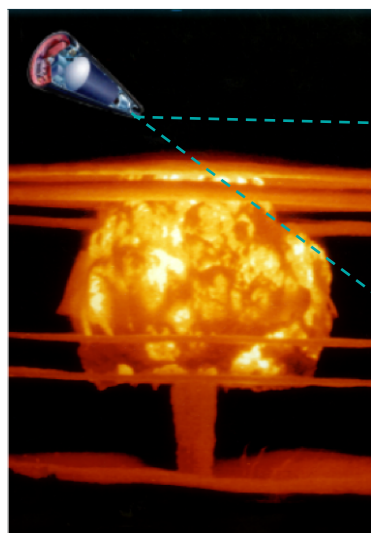**September 15, 2011**

Paul Lin
Sandia National Labs

SAND 2011-6489 P Unclassified Unlimited Release

Nuclear Survivability: Requirements to Certify Weapons System Electronics Within Hostile Environments
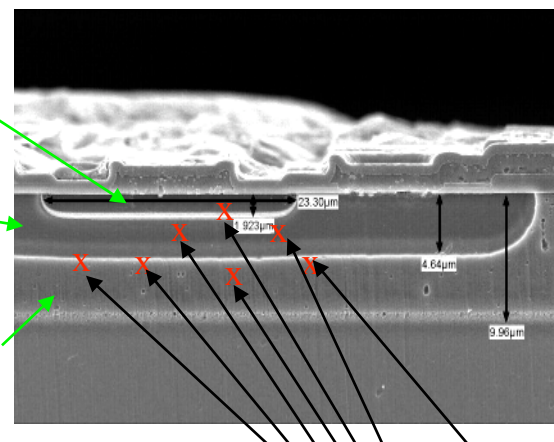
Neutrons create damage

Emitter (n-type)

Base (p-type)

Collector (n-type)

Damage degrades gain

Inverse Gain
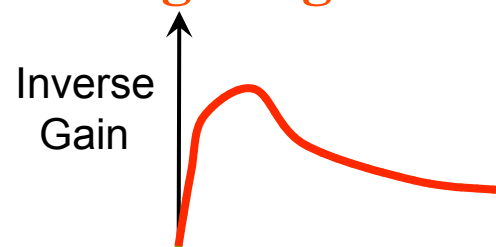
- **Sandia Pulse Reactor (SPR) testing provided qualification evidence in form of "Go/No-Go" decision**
- **SPR dismantled end of FY06**

QASPR (Qualification Alternatives to Sandia Pulse Reactor) will provide a methodology to provide evidence for qualification via quantified uncertainty

Slide courtesy of J. Castro

Sandia National Laboratories

RAMSES

Radiation Analysis, Modeling and Simulation for Electrical Systems

QASPR computational sequence

Neutron & gamma creation & propagation from reactor or nuclear burst
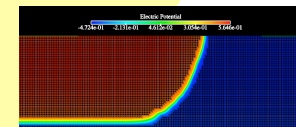**Neutron/Gamma Transport (NuGET)**
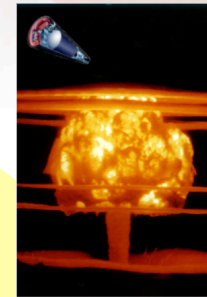
Defect recoil-cascade formation
Neutron Collision/Damage Formation (Cascade)
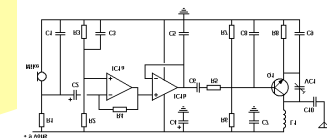(binary-collision approximation)
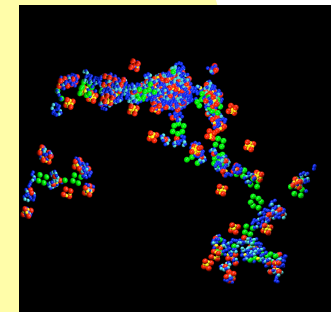
Defect & carrier reactions within recoil cascade
Defect Clustering (Cluster)

Finite-element device model with defect annealing
Device Performance (Charon)

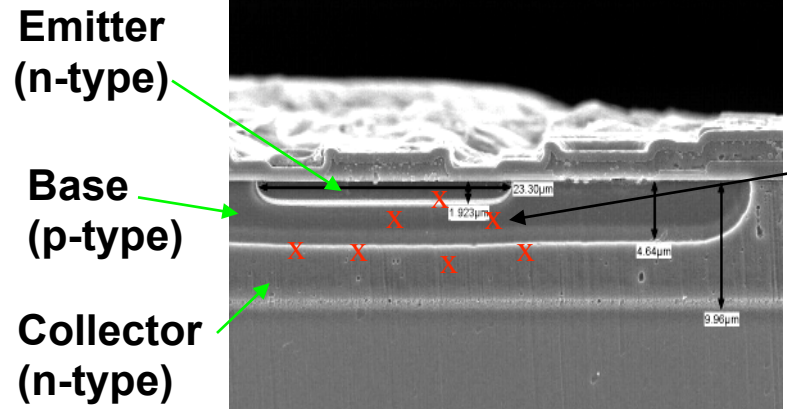Board & cable parasitics analysis
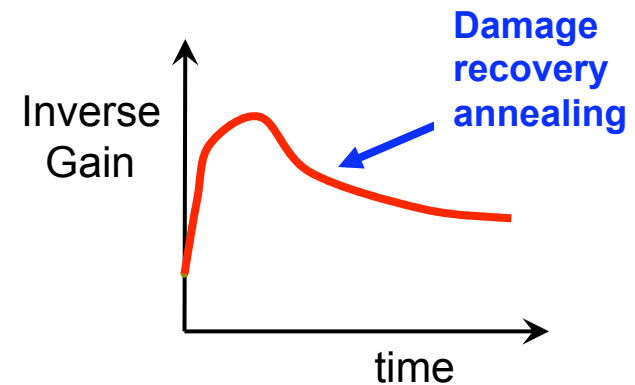
Spice model
Circuit Model (Xyce)

Slide courtesy J. Castro

Device Models Track the Transient Migration of Carriers and Defects Caused by Displacement Damage

Species & processes relevant for T = 300 K and times < 1 s.

Emitter (n-type)

Base (p-type)

Collector (n-type)

23.30μm
1.923μm
4.64μm
9.96μm

Displacement damage degrades device gain.

Si interstitial (I) (+2,+1,0,–1,–2)

B_I (+,0,–)

B_IB (+,0)

C_I (+,0,-)

B_IO (0,–)

Annihilation

VP (+,0,–)

VO (0,–)

VB (+,0)

Vacancy (V) (+2,+1,0,–1,–2)

VV (+1,0,–1,–2)

Divacancy

Carriers e⁻ & h⁺

33 species

~200 reactions
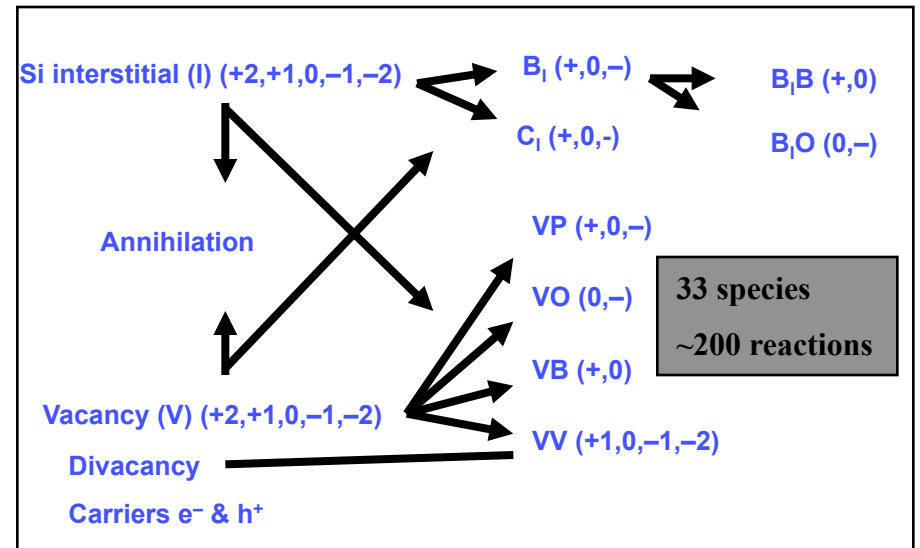
Defect Continuity	$$\frac{\partial Y_i(x)}{\partial t} + \nabla \cdot J_{Y_i} = R_{Y_i}(x)$$
Defect Species Flux	$$J_{Y_i} = \mu_i E Y_i(x) - D_i \nabla Y_i(x)$$
Poisson's Equation	$$-\nabla(\varepsilon \nabla \phi(x)) = -q(p(x) - n(x) + N_D^+(x) - N_A^-(x)) - \sum_{i=1}^{n} q_i Y_i(x)$$

Inverse Gain

Damage recovery annealing

time

UQ Process is Very Computationally Intensive
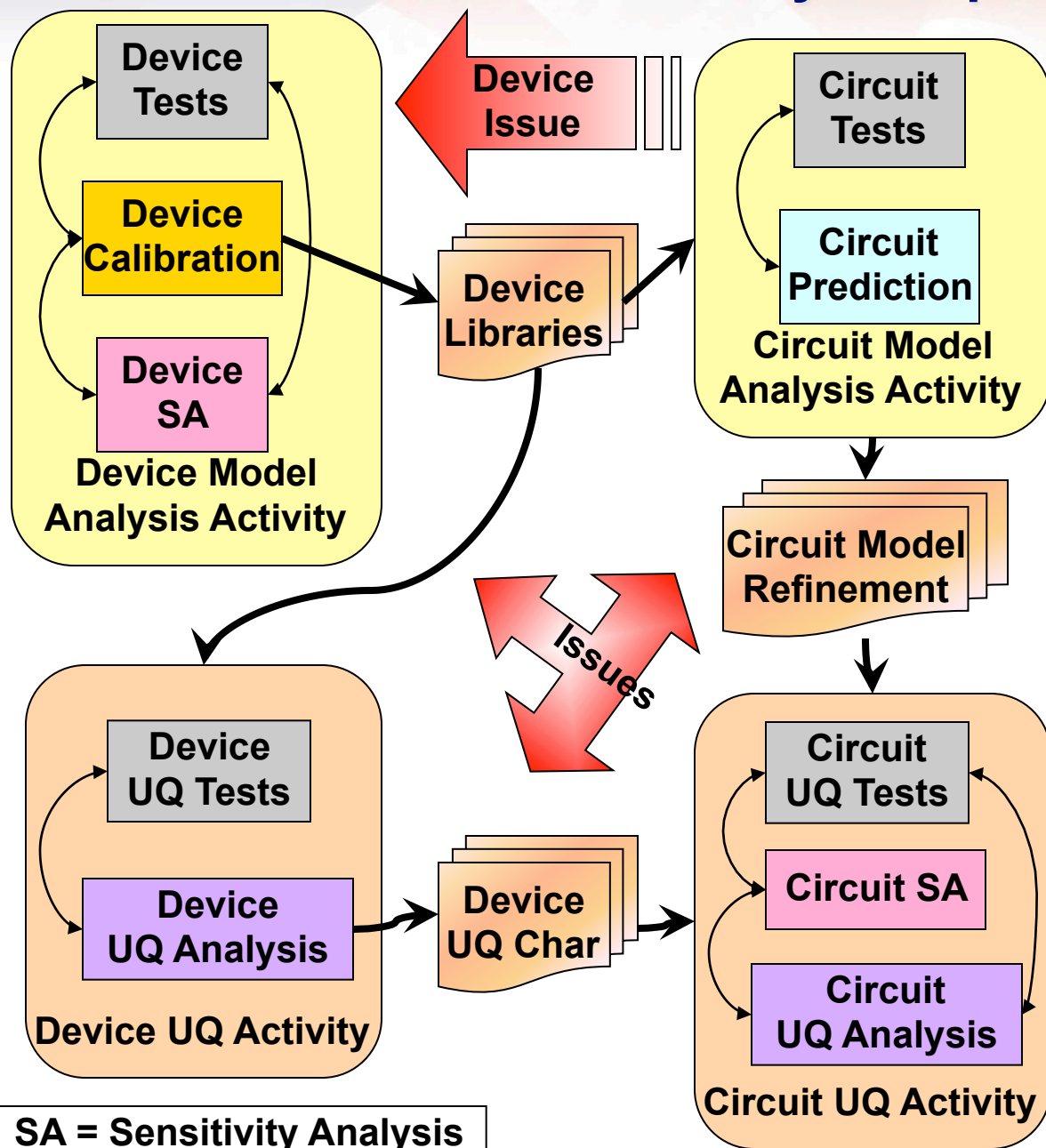
- UQ process includes: sensitivity analyses, calibration, and uncertainty propagation

- All these processes require ensembles of calculations

- UQ processes required at device and circuit level (very iterative process)

- Want to use highest fidelity models but limited by computational viability

SA = Sensitivity Analysis

Slide courtesy J. Castro

Sandia National Laboratories

# Computational Requirements for Semiconductor Device Simulations

- 2D bipolar junction transistor (BJT) with full defect physics $O(10^7 - 10^8)$ DOF; takes O(week) on $O(10^3)$ cores
- 3D simulations? $O(10^9 - 10^{10})$ unknowns
- Prediction plus uncertainly required for validation requires ensemble of calculations
  - 1D simulations presently (J. Castro et al.); $O(10^3)$ simulations
 - 2D could be performed on current largest platforms

1D
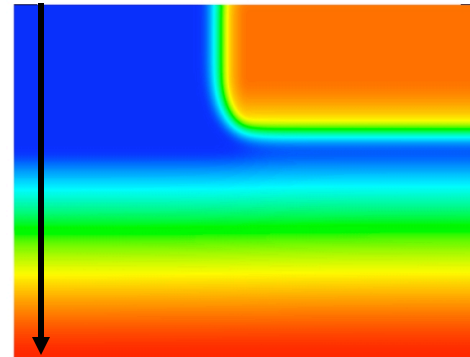
RAMSES/Charon Semiconductor Device Simulator

(Charon team: Hennigan, Hoekstra, Castro, Fixel, Pawlowski, Phipps, Musson, T. Smith, Shadid, Lin)

- Models the effects of radiation damage on semiconductor devices
- Drift-diffusion model; full defect physics for modeling damage to devices
- Massively parallel capability for high fidelity simulations
- FEM or FVM fully-implicit Newton-Krylov solver on unstructured meshes
- Fully-implicit Newton-Krylov robust; but need efficient solution of sparse linear systems
- SNL Trilinos solvers
  - Preconditioning
  - ML multigrid preconditioner
  - Currently using MPI-only portions of Trilinos

Sandia National Laboratories
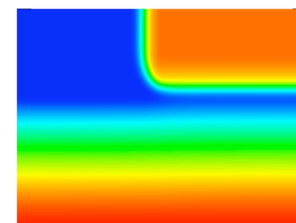
Preconditioners: Algebraic Multigrid for Semiconductor Problems

(w/ Shadid, Tuminaro, Hu, Siefert)

BG/P Weak Scaling 31,000 DOF/core
Charon Steady BJT GMRES

SNL Trilinos ML Library

(Tuminaro, Hu, Sala, Siefert, Tong, Gee)

- Algebraic multigrid has advantages over geometric multigrid for complex geometries
- Smoothed aggregation plus variant for nonsymmetric matrices

Level 2 (36 nodes)   Level 1 (9 nodes)   Level 0 (3) nodes
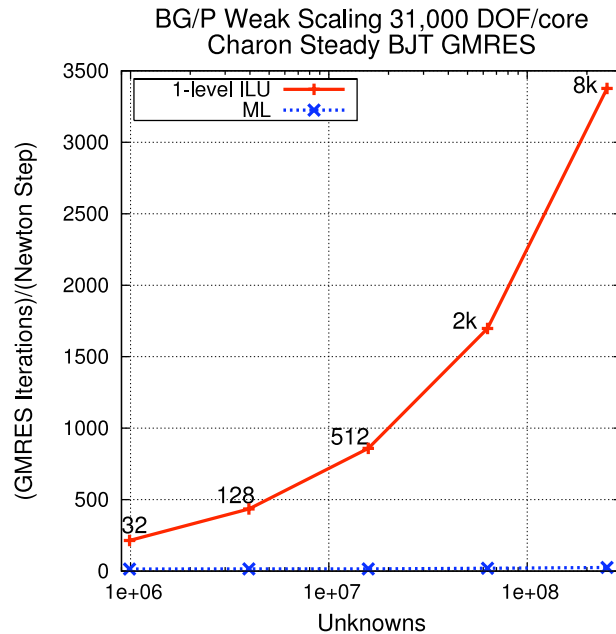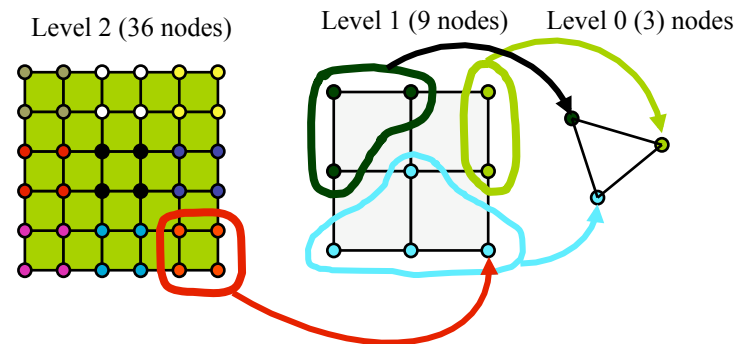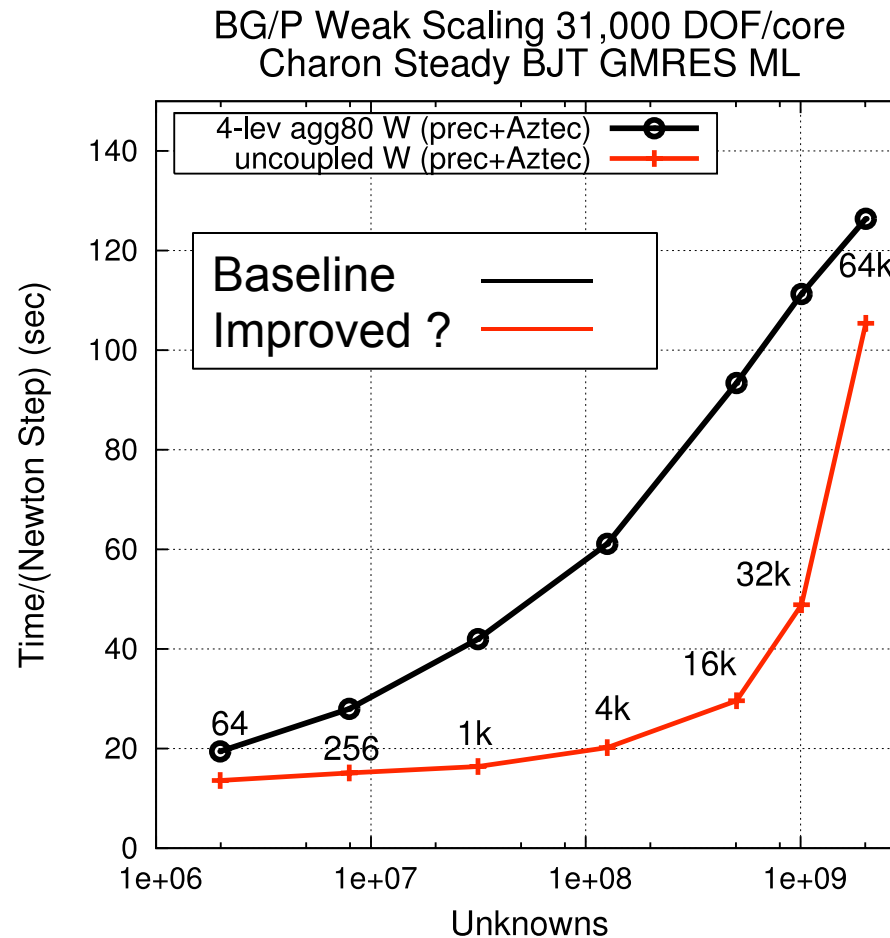
- Additive Schwarz domain decomposition preconditioners do not scale
- Need methods with global coupling such as multigrid

- Hypre (Falgout, Yang, Baker, Kolev, Tong, Chow, Lee, et al.) a very popular AMG library

Challenges of Debugging at Scale

BG/P Weak Scaling 31,000 DOF/core
Charon Steady BJT GMRES ML
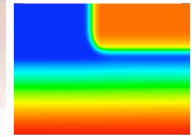
- Algorithm that works well on 4k cores may not work well for 64k cores
- Alternative promising algorithm scales well up to ~4k cores, but what happens?
- How to debug at 64k cores?
- Print statements ?!?!
- Need tools

- Steady drift-diffusion

Example Scalability Issue

BG/P Weak Scaling 31,000 DOF/core
Charon Steady BJT GMRES ML

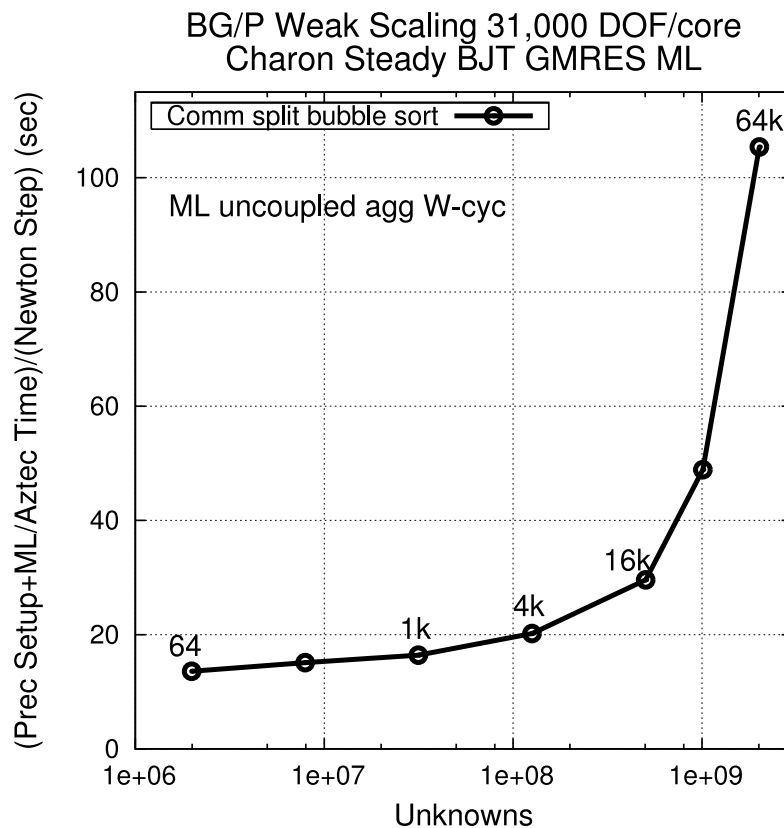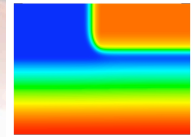Scaling of Comm_split for the case where the subcomm is the same as the communicator

cores	Split time (s)
8k	0.37
16k	1.6
32k	6.3
64k	28.2
128k	122
144k	154

- mpiP (Chambreau) indicated bottleneck was with Comm_split

Careful with MPI Implementation

Scaling of Comm_split for the case where the subcomm is the same as the communicator
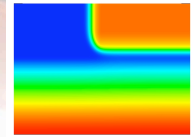
cores	bubble	quick
8k	0.37	0.11
16k	1.6	0.12
32k	6.3	0.14
64k	28.2	0.19
128k	122	0.29
144k	154	0.32

BG/P Weak Scaling 31,000 DOF/core
Charon Steady BJT GMRES ML

ML uncoupled agg W-cyc

Before fix ———
After fix ———

64k

64k

16k

4k

1k

64

(Prec Setup+ML/Aztec Time)/(Newton Step) (sec)

Comm split bubble sort ——○——
Comm split quick sort ——✕——

Unknowns
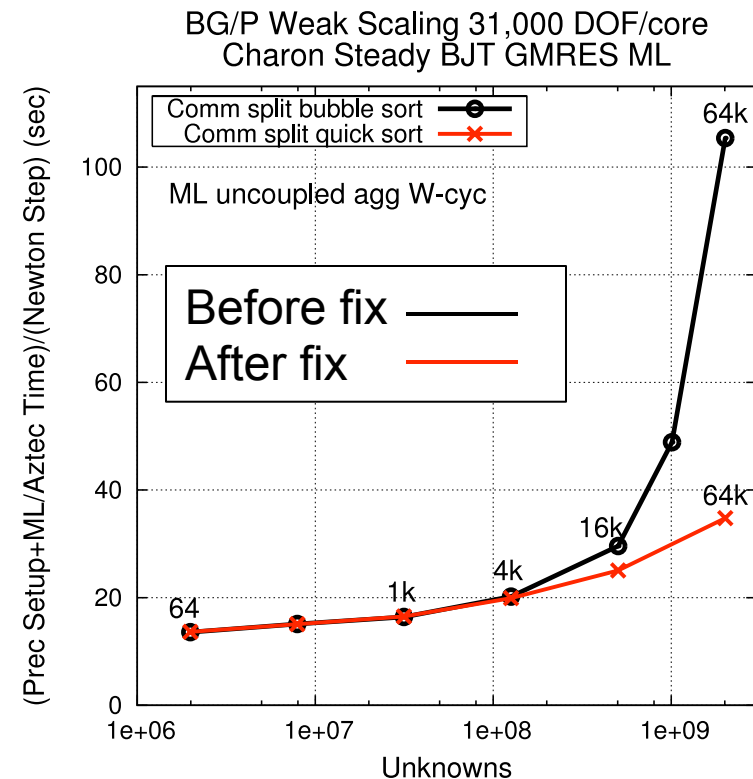
Message: Poor MPI implementation can hose a good algorithm

Message #2: When tracking down performance issues on O(10^5) processes, tool are critical

Improvements While Working on BG/P

BG/P Weak Scal 31,000 DOF/core: Charon Steady BJT GMRES ML
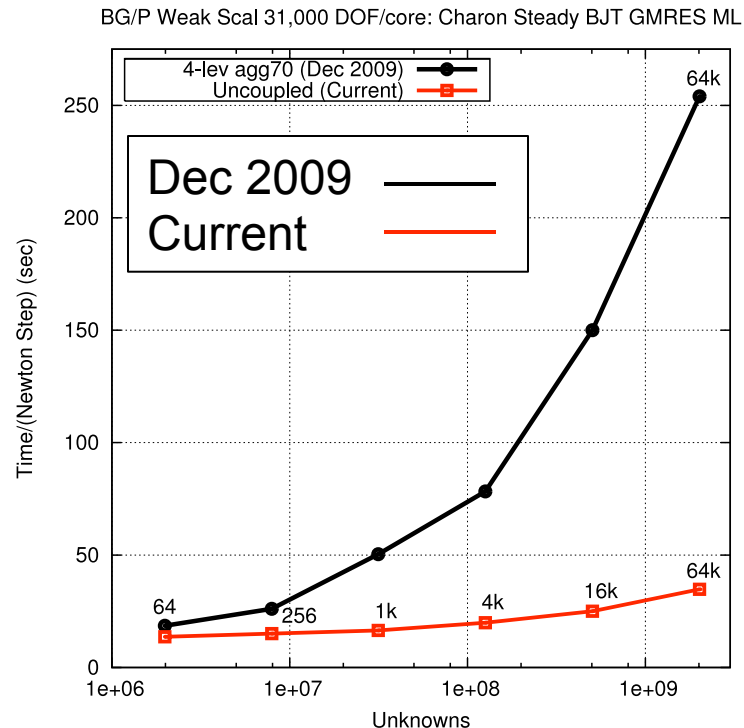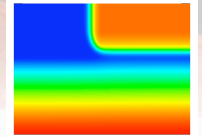
- Access to large number of cores critical to improving scalability
- BG/P: Reduced time by **7x** for 2 billion DOF on 64k cores compared with first scaling study (Dec 2009)
- Bigger machines important, but algorithmic improvements still critical

- Steady drift-diffusion

Improved Scalability… But More Work to Do

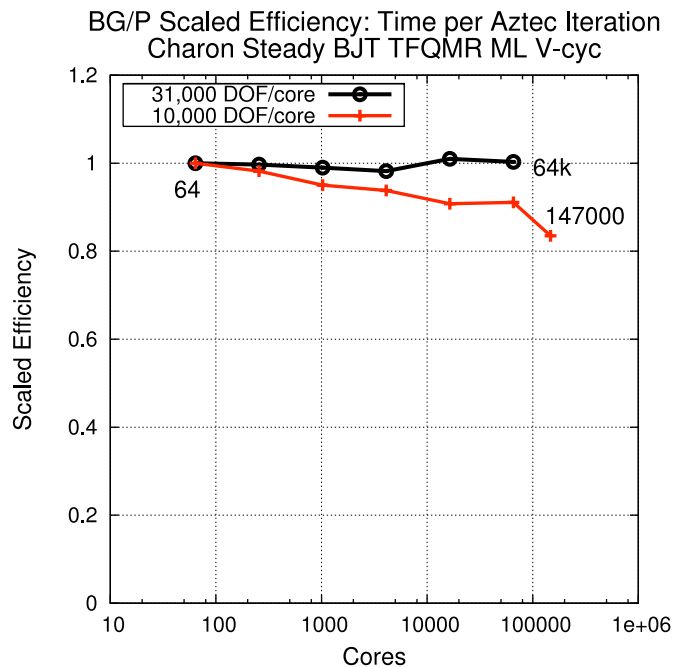- **Time per iteration scales well**
- But: preconditioner setup time, iteration count

BG/P Scaled Efficiency: Time per Aztec Iteration
Charon Steady BJT TFQMR ML V-cyc

- Close interaction with Trilinos team, so scaling improvements benefit other ASC codes
- Algorithms and scaling work directly benefited an SNL MHD project
- Hybrid MPI/threading/ accelerators: Trilinos Kokkos and "next generation" templated software stack (also fixes 32-bit global int problem)
- Muelu: next gen ML (Hu, Gaidamour, Tumi etc.)

- Scaled efficiency for time per iteration for 31,000 DOF/core (to 64k cores) and 10,000 DOF/core (to 147,000 cores)

Dawn BG/P Experience

- BG/P areas that could be improved
    - IBM C++ compiler buggy; poor templated code performance
    - Need alternative compiler, or GNU-built executable that worked
    - PowerPC slow, 1 GB RAM/core barely enough
    - Slow nodes (single bit errors); difficult when have 36k nodes
- BG/P has been a reliable platform
    - I/O always an issue for large machines, BG/P file system better than average
- BG/P an excellent resource to improve scaling
    - Platform has excellent scaling
    - Regular access to 64k cores (previously ~6k cores)
- Critical to have an unclassified porting/test platform
- Great support from LC
    - John G's help to port code (and feed bugs back to IBM)
    - Others (e.g. Scott, Tom, Sheila, etc.)

Concluding Remarks and Future Work

- Algorithmic improvements critical for scalability and efficiency
- Access to large core counts on Dawn critical for improving scalability of algorithms
- Tools critical (e.g. mpiP)
- Algorithmic improvements impacted Charon Cielo acceptance testing work
- Other codes (including ASC codes) will benefit from algorithmic improvements
- Murphy's Law exacerbated at scale: everything starts to break (app code, MPI implementation, tools, etc.)

Thanks For Your Attention!
Paul Lin (ptlin@sandia.gov)

Acknowledgments:

- ASC

- John Shadid

- Rest of Charon team: Gary Hennigan, Rob Hoekstra, Joe Castro, Roger Pawlowski, Debbie Fixel, Eric Phipps, Tom Smith, Larry Musson

- ML team: Ray Tuminaro, Jonathan Hu, Chris Siefert

- LLNL BG/P team
  - John Gyllenhaal: fighting with XL compiler bugs
  - Tom Spelce, Scott Futral, Adam Bertsch, Dave Fox, Chris Chambreau

- LLNL computing support: Sheila Faulkner, Tim Fahey

Thanks For Your Attention!

Paul Lin (ptlin@sandia.gov)

References:

- G Hennigan, R Hoekstra, J Castro, D Fixel, J Shadid, "Simulation of neutron radiation damage in silicon semiconductor devices," Technical Report SAND2007-7157, Sandia National Laboratories, 2007

- P Lin, J Shadid, M Sala, R Tuminaro, G Hennigan, R Hoekstra, "Performance of a Parallel Algebraic Multilevel Preconditioner for Stabilized Finite Element Semiconductor Device Modeling," Journ Comp Physics Vol 228 (2009), pp. 6250–6267

- P Lin and J Shadid, "Towards Large-Scale Multi-Socket, Multicore Parallel Simulations: Performance of an MPI-only Semiconductor Device Simulator," Journ Comp Physics Vol 229 (2010), pp. 6804–6818